U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

PB-291 842

Statistical Analysis of Factors Influencing Ozone Concentrations in Southern Louisiana

Radian Corp, Austin, TX

Prepared for

Environmental Protection Agency, Dallas, TX Region VI

Sep 78

PB 291842

EPA 906/9-78-002C

STATISTICAL ANALYSIS OF FACTORS INFLUENCING OZONE CONCENTRATIONS IN SOUTHERN LOUISIANA

Prepared by:

Hugh J. Williamson
Otto M. Friedrich

Radian Corporation
8500 Shoal Creek Boulevard
Austin, Texas 78766

Contract No. 68-02-1383, Task Order 18

TECHNICAL NOTE

Prepared for:

Franklin P. Hall, Jr.
Project Officer (6ASAS)
United States Environmental Protection Agency
First International Building
1201 Elm Street
Dallas, Texas 75270

September 1978

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

DISCLAIMER

This report was furnished to the Environmental Protection Agency by Radian Corporation, Austin, Texas, in fulfillment of Contract No. 68-02-1383, Task Order 18. This report has been reviewed by Region VI, EPA, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Region VI Publication No. 906/9-78-002C.

ACKNOWLEDGEMENTS

The authors appreciate the direction of Dr. Joe D. Stuart of Radian Corporation during the course of this study. Consultations with Mr. Bryan Lambeth and Mr. Norris Nielsen of Radian have also been very helpful. Finally, the assistance in data preparation by Mr. Mike A. Haecker was essential to the study.

TABLE OF CONTENTS

		Page
1.0	INTRODUCTION----------	1
2.0	CORRELATIONS OF OZONE CONCENTRATION WITH METEOROLOGICAL AND SPATIAL VARIABLES AND WITH CONCENTRATIONS OF OTHER POLLUTANTS---------	2
	2.1 Meaning of the Correlation Coefficient-----	2
	2.2 Correlations of Ozone with Meteorological and Spatial Variables----------	9
	2.3 Correlations of Ozone with Other Pollutants----------	13
	2.4 Regression Model for Average Ozone Concentrations----------	13
3.0	EFFECT OF LOCAL TRANSPORT----------	17
	3.1 Local Transport Interaction Terms----------	17
	3.2 Effects of Wind Direction at Individual Monitoring Sites----------	21
4.0	DEPENDENCIES AMONG THE METEOROLOGICAL AND SPATIAL VARIABLES----------	26
5.0	CORRELATION OF SOLAR RADIATION WITH OTHER VARIABLES----------	50
6.0	OZONE CORRELATIONS AMONG SITES----------	55
7.0	SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS-------	60
	APPENDIX A - ADDITIONAL INFORMATION REGARDING THE MULTIPLE REGRESSION MODEL FOR AVERAGE OZONE CONCENTRATION----	65

LIST OF FIGURES

		Page
2-1	Positions of Monitoring Sites in Southern Louisiana----------	6
2-2	Illustration of the Meaning of Correlation, r---	8
2-3	Hydrocarbon and Oxides of Nitrogen Emission Densities----------	12

LIST OF TABLES

		Page
2-1	Definition of Variables----------	3
2-2	Correlations of Average Ozone Concentration with Meteorological and Spatial Var. Tables----------	10
2-3	Correlations Between Average Concentrations of Different Pollutants----------	14
2-4	Correlations Between Maximum Concentrations of Different Pollutants----------	14
3-1	Correlation of Pollutant Concentrations with Wind-Direction Variables----------	20
3-2	Correlation of Pollutant Concentrations with First Set of Local Transport Terms----------	20
3-3	Correlation of Pollutant Concentrations with Second Set of Local Transport Terms----------	20
3-4	Site-Specific Correlations of Average Ozone Concentration with Wind Direction----------	22
3-5	Regression Equations for Average Ozone Concentration at Site 8----------	24
3-6	Comparison of Regression Equations With and Without the 16 Wind-Direction Terms----------	25
4-1	Correlations Significant at 90.0 PCT. Level Between WS AV and Other Variables----------	28

LIST OF TABLES (contd) Page

4-2	Correlations Significant at 90.0 PCT. Level Between TEMPAV and Other Variables----------	29
4-3	Correlations Significant at 90.0 PCT. Level Between WD SD and Other Variables----------	30
4-4	Correlations Significant at 90.0 PCT. Level Between DP AV and Other Variables----------	31
4-5	Correlations Significant at 90.0 PCT. Level Between WD HOU and Other Variables----------	32
4-6	Correlations Significant at 90.0 PCT. Level Between TEMPMX and Other Variables----------	33
4-7	Correlations Significant at 90.0 PCT. Level Between TEMP12 and Other Variables----------	34
4-8	Correlations Significant at 90.0 PCT. Level Between Site 8 and Other Variables----------	36
4-9	Correlations Significant at 90.0 PCT. Level Between Site 10 and Other Variables----------	37
4-10	Correlations Significant at 90.0 PCT. Level Between Site 12 and Other Variables----------	38
4-11	Correlations Significant at 90.0 PCT Level Between Site 16 and Other Variables----------	39
4-12	Correlations Significant at 90.0 PCT Level Between WD AV and Other Variables----------	40
4-13	Correlations Significant at 90.0 PCT Level Between TEMPAV and Other Variables----------	41

LIST OF TABLES (contd)

		Page
4-14	Correlations Significant at 90.0 PCT. Level Between WD SD and Other Variables----------	42
4-15	Correlations Significant at 90.0 PCT. Level Between DP AV and Other Variables----------	43
4-16	Correlations Significant at 90.0 PCT. Level Between WD HOU and Other Variables----------	44
4-17	Correlations Significant at 90.0 PCT. Level Between TEMPMX and Other Variables----------	45
4-18	Correlations Significant at 90.0 PCT. Level Between WDIR11 and Other Variables----------	46
4-19	Correlations Significant at 90.0 PCT. Level Between WDIR15 and Other Variables----------	47
4-20	Correlations Significant at 90.0 PCT. Level Between TEMP12 and Other Variables----------	48
5-1	Correlations of Several Variables With Average Ozone Concentration----------	51
5-2	Partial Correlations of Solar Radiation With Average Ozone Concentration----------	53
5-3	Multiple Correlations of Sets of Variables With Average Ozone Concentration----------	53
6-1	Ozone Means and Standard Deviations by Site for 42 Hours Occurring on September 5-9, 1976-------	56
6-2	Site-by-Site Correlations Between Hourly Ozone Concentrations, Ordered by Site Location--------	57
6-3	Site-by-Site Correlations Between Hourly Ozone Concentrations, Ordered by County NO_x and Hydrocarbon Emission Densities----------	58

EXECUTIVE SUMMARY

This report discusses a statistical study of ozone concentrations in southern Louisiana from July 17 to October 17, 1976. Data collected at nine monitoring sites were analyzed.

It has been shown that ozone concentration increases as the following variables increase:

(1) wind direction variability,
(2) ambient temperature, and
(3) solar radiation.

Ozone concentration decreases as the following variables increase:

(1) average wind speed and
(2) average dew point.

Of these five, solar radiation has the highest correlation, .392, with average ozone concentration. As is indicated below, it is believed that dew point and ambient temperature correlate with ozone at least partly because of their correlation with solar radiation.

There is not an extremely strong area-wide dependence of ozone concentration on wind direction. Ozone tends to be slightly lower, however, when the wind is from the west and northwest. There is no evidence that transport from the New Orleans or Houston-Port Arthur areas increased the ozone concentrations at the monitoring sites. There is evidence, however, that transport from Baton Rouge increased ozone concentrations to a small extent.

It is felt that the following are important direct influences on ozone concentration.

(1) Strong winds disperse pollutants of all kinds, and, therefore, reduce their concentrations.

(2) Vertical mixing tends to reduce the concentrations of most pollutants but can increase the ozone concentration in the morning by bringing ozone from aloft to ground level. Vertical mixing is greatest when atmospheric instability exists, which is most likely when wind speeds are low and wind speed variability is large.

(3) Solar radiation promotes ozone formation through atmospheric photochemical reactions. Solar radiation has been shown to have a strong statistical relationship with ambient temperature and dew point.

1.0 INTRODUCTION

This report describes a statistical analysis of ozone concentrations in southern Louisiana. The data which have been used were collected at nine monitoring sites during a three-month period, July 17 - October 17, in 1976. The positions of the monitoring sites are shown in Figure 2-1, which is presented in the next section.

The objective of the study was to investigate the factors which accompany high levels of ozone concentration and to explain the observed relationships to as great an extent as possible. This was done to gain a better understanding of the variables which influence the amount of ozone. The relationships between ozone concentration and a number of meteorological and spatial variables have been explored. The relationships between ozone concentration and the concentrations of other pollutants are also discussed briefly.

2.0 CORRELATIONS OF OZONE CONCENTRATION WITH METEOROLOGICAL AND SPATIAL VARIABLES AND WITH CONCENTRATIONS OF OTHER POLLUTANTS

The purpose of this section is to discuss the relationships of ozone with a number of meteorological and spatial variables and with the concentrations of other pollutants. The variables which were investigated are listed and defined in Table 2-1.

As is indicated by the table, the concentrations of pollutants were averaged from 10 a.m. to 4 p.m. for each day for each site. The mid-day period was of most interest, since this is when the ozone concentration is generally highest. The meteorological variables, which relate to ozone formation and transport, were averaged from 8 a.m. until 4 p.m. A similar scheme was used in a study of ozone concentrations in Los Angeles and St. Louis, which is discussed in Reference 2. As Table 2-1 indicates, the maximum daily values of the pollutant concentrations and of ambient temperature were also used. Solar radiation, which was measured at only three sites, is discussed in a subsequent section.

The positions of the monitoring sites within southern Louisiana are shown in Figure 2-1.

2.1 Meaning of the Correlation Coefficient

The correlation coefficient is a very useful measure of the extent to which two variables are related. For this reason, correlation has been used in this study to quantify relationships between pairs of variables.

TABLE 2-1. Definition of Variables

Variable Number	Variable Acronym	
	Variables 1 - 9 are "dummy" variables which indicate the monitoring site being considered. Variable 1, for example, is 1 if the data being considered were measured at Site 8 and is 0 otherwise.	
1	SITE 8	
2	SITE10	
3	SITE12	
4	SITE13	
5	SITE14	
6	SITE16	
7	SITE17	
8	SITE18	
9	SITE19	
	Variables 10 - 14 represent average pollutant concentrations in ppb from 10 a.m. to 4 p.m. at a particular monitoring site.	
10	O_3 AV	average ozone concentration
11	NO_x AV	average NO_x concentration
12	NO AV	average NO concentration
13	NO_2 AV	average NO_2 concentration
14	SO_2 AV	average SO_2 concentration
	Variables 15 - 21 represent conditions from 8 a.m. to 4 p.m. at a particular monitoring site.	

TABLE 2-1 (cont'd)

Variable Number	Variable Acronym	
15	WS AV	average wind speed in miles per hour
16	TEMPAV	average ambient temperature in ^{o}F
17	WD SD	wind direction standard deviation in degrees*
18	DP AV	average dew point in ^{o}F
19	WD NOR	number of hours that the average wind direction was toward the monitoring site from New Orleans
20	WD BTR	same as variable 19, except from Baton Rouge
21	WD HOU	same as variable 19, except from Houston
	Variables 22 - 26 are maximum pollutant concentrations in ppb between 10 a.m. and 4 p.m. at a given site.	
22	O_3 MX	maximum ozone concentration
23	NO_x MX	maximum NO_x concentration
24	NO MX	maximum NO concentration
25	NO_2 MX	maximum NO_2 concentration
26	SO_2 MX	maximum SO_2 concentration
27	TEMPMX	maximum ambient temperature between 8 a.m. and 4 p.m. at a given site
28	D NOR	distance in nautical miles from a monitoring site to New Orleans
29	D BTR	distance in nautical miles from a monitoring site to Baton Rouge
30	D HOU	distance in nautical miles from a monitoring site to the center of the Houston-Port Arthur area.

TABLE 2-1 (cont'd)

Variable Number	Variable Acronym		
		Variables 31 - 46 indicate the number of hours between 8 a.m. and 4 p.m. that the wind was from a given direction. Zero degrees is north, and the angles are positive in the clockwise sense.	
31	WDIR 1	0-22.5°	north
32	WDIR 2	22.5-45°	
33	WDIR 3	45-67.5°	
34	WDIR 4	67.5-90°	east
35	WDIR 5	90-112.5°	
36	WDIR 6	112.5-135°	
37	WDIR 7	135-157.5°	
38	WDIR 8	157.5-180°	south
39	WDIR 9	180-202.5°	
40	WDIR10	202.5-225°	
41	WDIR11	225-247.5°	
42	WDIR12	247.5-270°	west
43	WDIR13	270-292.5°	
44	WDIR14	292.5-315°	
45	WDIR15	315-337.5°	
46	WDIR16	337.5-360°	north
47	TEMP12	average ambient temperature between noon and 1 p.m.	

*For a given hour, the standard deviation of 300 one-second wind directions was computed for each successive five-minute period. The twelve resulting values were then averaged to obtain WD SD. Thus, WD SD is not, strictly speaking, a standard deviation.

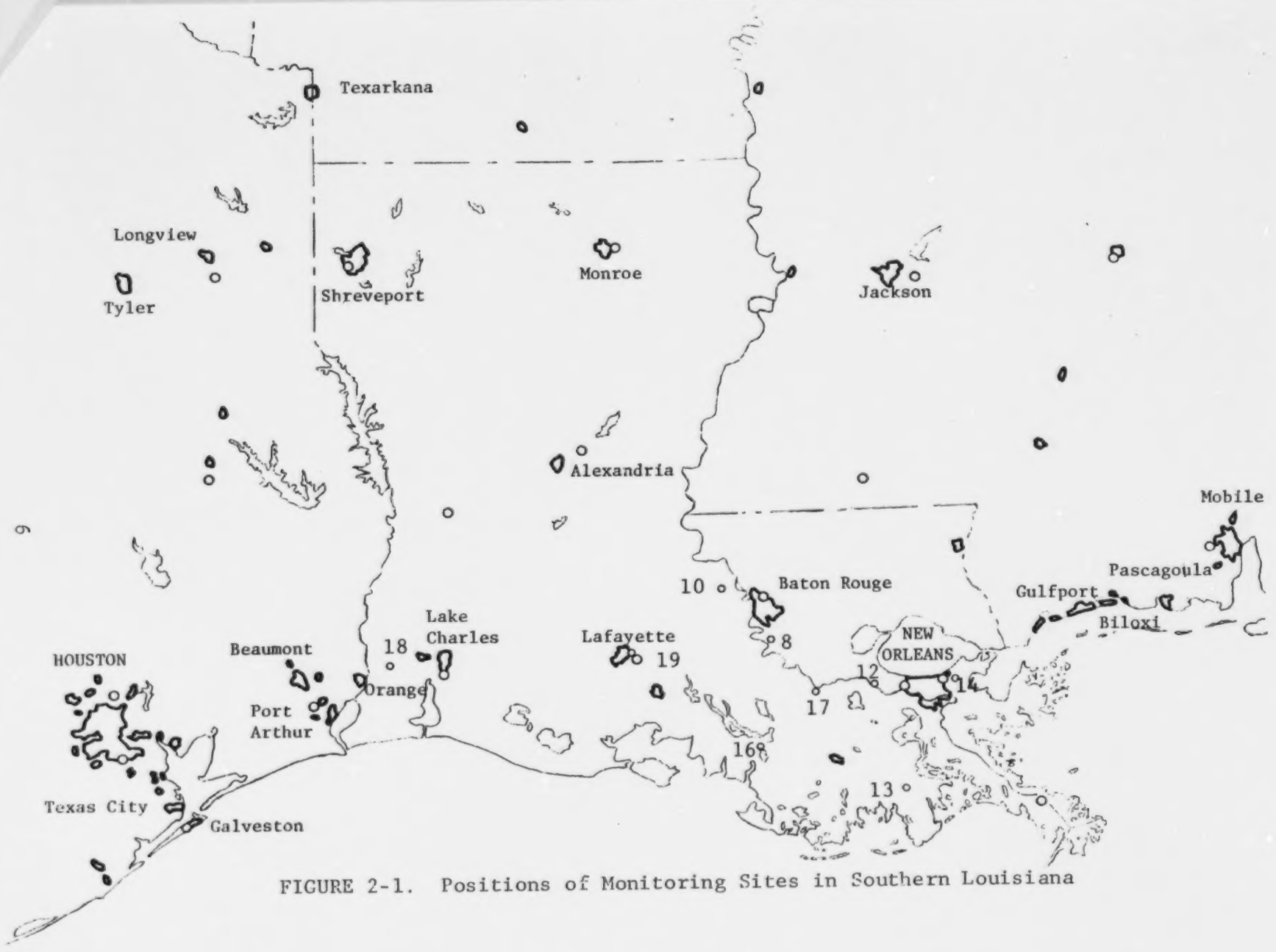


FIGURE 2-1. Positions of Monitoring Sites in Southern Louisiana

The correlation coefficient r is defined as follows:

$$r = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2 \sum_{j=1}^{N} (y_j - \bar{y})^2}}$$

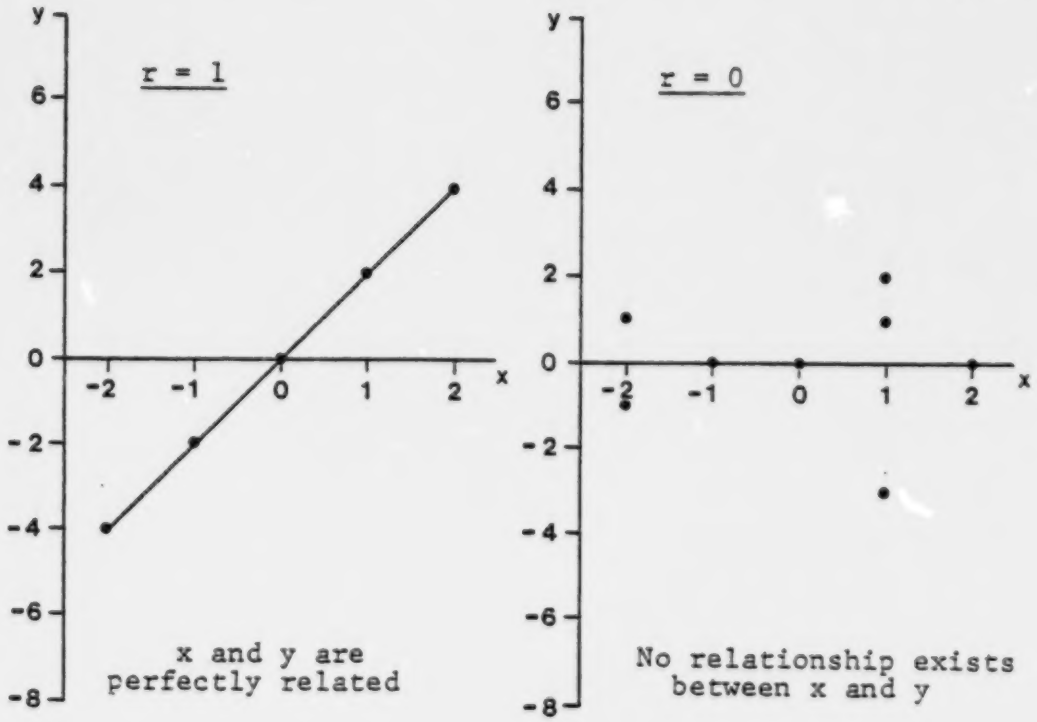
where

N = number of data points,
$\bar{x}$ = average of the x values, and
$\bar{y}$ = average of the y values.

The correlation is a number between -1 and +1, and r^2 can be interpreted as the proportion of the variation in y that can be "explained" or "predicted" as a linear function of x. If, for example, the correlation were .7, about half of the variation in y could be explained in terms of x.

Figure 2-2 illustrates the meaning of a non-negative correlation coefficient. A negative correlation between two variables occurs if one increases as the other decreases.

The correlation coefficient is defined on p. 33 of Reference 1 as a measure of the relationship between two "random variables following some (unknown) bivariate distribution...." Use of the correlation coefficient does not require that the two variables have the normal (or any other specific) distribution. Correlation is used with discrete as well as continuous random variables.



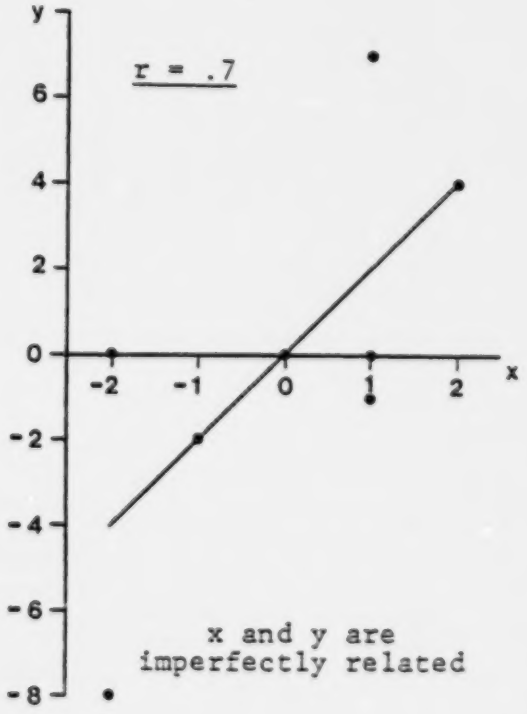


FIGURE 2-2. Illustration of the Meaning of Correlation, r

2.2 Correlations of Ozone with Meteorological and Spatial Variables

Table 2-2 presents the mean, standard deviation, and correlation with the average ozone concentration of various meteorological and spatial variables.

The correlations which are statistically significant at the 90 percent level are marked with asterisks. In these cases, there is strong reason to think that a "true" or "measurable" relationship exists between the indicated variables and the average ozone concentration. Each of the significant correlations exceeds .1 in magnitude; there is only one chance in ten that this would occur if a variable, in fact, had no actual correspondence with the ozone concentration. The threshold value, .1, is dependent on the data sample size, which is 273 in this case.

It should be noted that, although several of the variables are measurably correlated with ozone concentration, none of the relationships are close. The correlations which do exist, however, generally have plausible explanations, as is discussed below.

First, the average temperature, maximum temperature, and temperature at noon all correlate positively and significantly with ozone concentration. This is to a large extent because these variables are related to solar radiation, which promotes ozone formation. According to Reference 6, ozone formation in the lower atmosphere "usually is attributed to the nitrogen dioxide photolytic cycle":

TABLE 2-2

CORRELATIONS OF AVERAGE OZONE CONCENTRATION WITH METEOROLOGICAL AND SPATIAL VAR.

VARIABLE	MEAN	STD. DEV.	CORRELATION WITH O_3 AV.
SITE 9	.14	.35	.243*
SITE10	.18	.38	-.118*
SITE12	.10	.30	-.190*
SITE13	.05	.23	.053
SITE14	.13	.34	.082
SITE16	.10	.29	-1.05*
SITE17	.01	.12	.024
SITE18	.10	.29	-.088
SITE19	.18	.39	.082
O3 AV	65.17	25.73	1.000*
WS AV	4.20	2.04	-.198*
TEMPAV	80.94	5.44	.163*
WD SD	13.66	3.12	.159*
DP AV	69.17	6.65	-.261*
WD NOR	.49	.89	.027
WD BTR	.42	.81	.039
WD HOU	.58	1.04	-.102*
TEMPMX	85.44	5.57	.223*
D NOR	68.99	47.08	-.056
D BTR	47.44	34.03	-.082
D HOU	178.10	60.93	.030
WDIR 1	.13	.48	-.008
WDIR 2	.60	1.19	.022
WDIR 3	.82	1.31	.020
WDIR 4	.78	1.21	.057
WDIR 5	.79	1.13	.099
WDIR 6	.63	.89	.065
WDIR 7	.49	.89	.047
WDIR 8	.44	.73	-.052
WDIR 9	.40	.82	.038
WDIR10	.39	.74	.027
WDIR11	.45	.82	-.149*
WDIR12	.52	.98	-.099
WDIR13	.44	.82	-.049
WDIR14	.59	1.16	.028
WDIR15	.42	1.03	-.102*
WDIR16	.10	.42	-.026
TEMP12	83.01	5.78	.252*

*Correlation is statistically significant at the 90 percent level.

$$NO_2 + (photon) \rightarrow NO + O$$
$$O + O_2 + M \rightarrow O_3 + M$$
$$O_3 + NO \rightarrow NO_2 + O_2$$

where M is any third body which absorbs the excess energy of the initially formed O_3 molecule and, thus, allows it to be stabilized.

Although many other atmospheric chemical reactions involving ozone, NO_x, and hydrocarbons also occur, the effect of the photochemical reactions is believed to be a net production of ozone.

Second, ozone is negatively correlated with wind speed. This is probably because high winds tend to disperse pollutants from areas with high pollution emissions, and the monitoring sites are all in or near parishes with high hydrocarbon and NO_x emission densities. Figure 2-3 displays the emission densities by parish in the area. This figure was taken from Reference 3, with permission of the author.

The wind direction standard deviation (WD SD) correlates positively with ozone concentration. There are two possible explanations for this. First, large values of WD SD are associated with atmospheric instability, which is associated with vertical mixing. There is evidence in the literature, moreover, that large concentrations of ozone, in excess of 20 pphm, exist overnight within elevated inversions (Reference 5). Thus, vertical mixing in the morning can contribute to ground-level ozone by bringing the ozone aloft to lower levels.

A second possible explanation is that variable winds are generally lighter and, thus, less dispersing. Wind speed and WD SD have a correlation of -.444. The relationships among WD SD, WS AV, O_3 AV, and atmospheric stability are discussed further in Section 4.0.

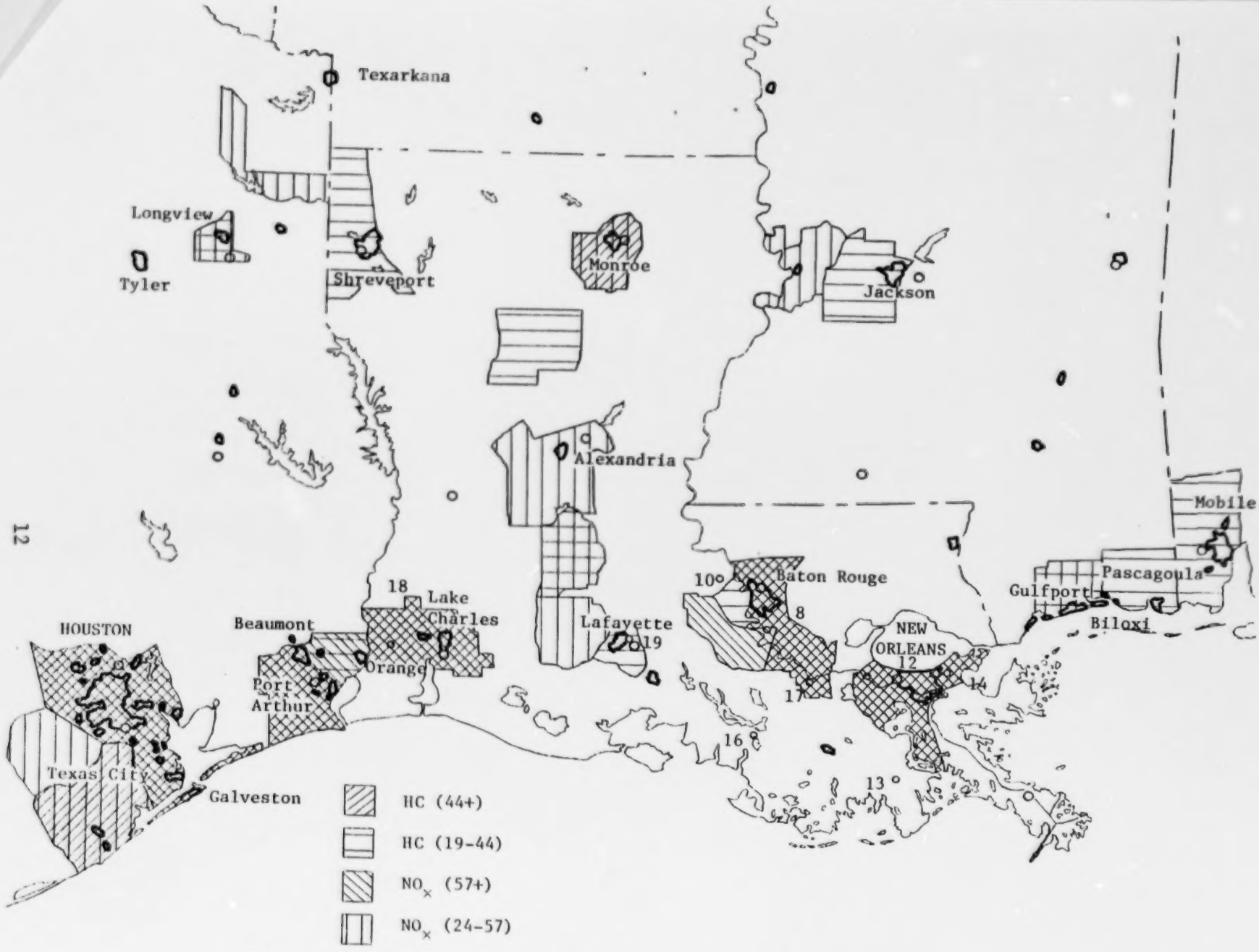


FIGURE 2-3. HYDROCARBON AND OXIDES OF NITROGEN EMISSION DENSITIES (tons/sq.mi./yr.)

Other relationships among the variables are also discussed in subsequent sections in connection with additional statistical analyses.

2.3 Correlations of Ozone with Other Pollutants

Table 2-3 presents the correlations among the average pollution concentrations, and Table 2-4 presents the correlations among the maximum concentrations. It can be seen that, although NO, NO_2, NO_x, and SO_2 all correlate positively, the correlations between ozone and the other pollutants are all very low.

The correlation between the average and the maximum ozone concentrations is .942, which is very high.

2.4 Regression Model for Average Ozone Concentration

The following regression equation was developed for average ozone concentration:

$(O_3$ AV) =
24.4 + 17.8 (SITE 8) + 21.4 (SITE 13) + 19.8 (SITE 14)
+ 24.7 (SITE17) + 20.8 (SITE19)
- 3.34 (WS AV) - 2.60 (DP AV)
+ 0.0542 (D HOU) + 1.98 (WDIR 2)
- 2.30 (WDIR 8) + 2.46 (WDIR10)
- 6.02 (WDIR11) + 2.89 (WDIR14)
+ 2.58 (TEMP12)

The correlation coefficient for the equation is .732, and the standard error is 18.0 ppb. The standard error can be thought of as the standard deviation of the differences between the predicted and measured values. The equation was developed from 273 data points.

TABLE 2-3. Correlations Between Average Concentrations of Different Pollutants

	O_3 AV	NO_x AV	NO AV	NO_2 AV	SO_2 AV
O_3 AV	1.				
NO_x AV	-.039	1.			
NO AV	-.090	.683	1.		
NO_2 AV	-.005	.928	.366	1.	
SO_2 AV	-.049	.274	.397	.153	1.

TABLE 2-4. Correlations Between Maximum Concentrations of Different Pollutants

	O_3 MX	NO_x MX	NO MX	NO_2 MX	SO_2 MX
O_3 MX	1.				
NO_x MX	.025	1.			
NO MX	-.044	.830	1.		
NO_2 MX	.073	.866	.442	1.	
SO_2 MX	-.029	.314	.353	.190	1.

Stepwise regression analysis (Reference 1) was employed to develop the equation. This method is used to select from a set of candidate predictor variables only those which contribute significantly to the accuracy of the equation. All the spatial and meteorological variables listed in Table 2-1 were included as candidates. Thus, the correlation and standard error given above can be thought of as indicating the composite relationship between these variables and the average ozone concentration.

The inclusion of terms in the model is a function of their interdependence, as well as their isolated relationships with ozone. Of the three temperature variables, TEMPAV, TEMPMX, and TEMP12, for example, only one, TEMP12, was included in the model. This is because of the high correlations among the three variables. Given that TEMP12 is in the model, the other two variables make no further significant contribution to the predictive accuracy.

Of the nine site variables (SITE 8, SITE 10, etc.), five were selected for inclusion in the model by the stepwise algorithm. This does not mean that the other four effects are identically zero; as with all other excluded variables, the other four site variables would have made no further statistically significant contribution to the predictive accuracy.

Various plots of measured minus predicted ozone concentrations versus predictor variables were produced, and no evidence was seen that the inclusion of a nonlinear function of any one of the variables would improve the equation.

Various interaction terms involving WD BTR, WD HOU, and WD NOR with certain other variables were considered, and none contributed to the predictive accuracy. This point is discussed further in the section on local transport.

The assumptions made in the regression analysis and additional statistical details are discussed in Appendix A.

It is mentioned above that correlations among the independent variables affect which variables are selected for inclusion in the model. Correlations among variables which are included, moreover, affect the values of the regression coefficients. These effects are such that interpretation of multiple regression coefficients in many cases is extremely difficult.

For these reasons, the meaning of the regression coefficients is not discussed here; the relationships between individual independent variables and ozone concentration are instead discussed in connection with the correlation and partial correlation analyses presented in Sections. 2.2, 3, 4, and 5.

These comments are not meant to imply that the regression analysis is not useful. As is discussed above, it provides measures of the closeness of the relationship between (1) average ozone concentration and (2) the spatial and meteorological variables as a group.

3.0 EFFECT OF LOCAL TRANSPORT

In this section, the effect of local transport on ozone concentrations is investigated.

3.1 Local Transport Interaction Terms

As is indicated by Figure 2-3, there are hydrocarbon and NO_x sources in the Baton Rouge, New Orleans, and Houston-Port Arthur areas. Shown as emission densities by parish, the emissions are actually concentrated in the urban areas. Lake Charles is a small urban area with smaller emissions than the cities mentioned above.

The contribution of transport from these areas was investigated by studying the correlations between average ozone concentration and variables WD NOR, WD BTR, and WD HOU. As defined in Table 2-1, these three variables refer to the number of hours from 8 a.m. to 4 p.m. that the wind was directed toward a given monitoring site from New Orleans, Baton Rouge, or Houston, respectively.

These variables by themselves are inadequate, since they do not reflect (1) the different distances from the various sites to a given metropolitan area or (2) the variation in the amount of ozone formed on a given day. Transport is most important when the pollutant to be transported is in large concentration at the source.

To account for (1), the WD variables were divided in each case by the square of the distance from the monitoring site to the urban area. To account for (2), this quotient was multiplied by the average temperature at the monitoring site from 8 a.m. to 4 p.m.; due to the relatively small geographical

area involved, it is felt that the temperature at the monitoring sites and in the urban areas are closely related.

Using the symbolism of Table 2-1, then, the resulting interaction terms are as follows:

$$T_{NOR} = \frac{(WD\ NOR)(TEMPAV)}{(D\ NOR)^2}$$

for New Orleans,

$$T_{BTR} = \frac{(WD\ BTR)(TEMPAV)}{(D\ BTR)^2}$$

for Baton Rouge,

$$T_{HOU} = \frac{(WD\ HOU)(TEMPAV)}{(D\ HOU)^2}$$

for Houston-Port Arthur.

Wind speed is also a factor in transport. In the general case, strong winds tend to disperse pollutants and, therefore, to decrease their concentrations. High wind speed, however, is also associated with wind persistence; recall that average wind speed and the wind direction standard deviation have a correlation of -.444. One would expect that transport would be most effective when a persistent wind were blowing from a source toward a monitoring site.

The three following transport terms involving wind speed, then, were also considered.

$$T'_{NOR} = (T_{NOR})(WS\ AV)$$
$$T'_{BTR} = (T_{BTR})(WS\ AV)$$
$$T'_{HOU} = (T_{HOU})(WS\ AV)$$

Table 3-1 presents the correlations of the pollutant concentrations with the WD variables; Table 3-2, with T_{NOR}, T_{BTR}, and T_{HOU}; and Table 3-3, with T^1_{NOR}, T^1_{BTR}, and T^1_{HOU}.

It is clear that ozone concentration does not correlate highly with any of the local transport variables. There are, however, significant correlations between T^1_{NOR} and concentrations of the oxides of nitrogen. New Orleans is by far the largest urban area which is close to most of the monitoring sites. Importantly, the correlation of T^1_{NOR} with NO is almost twice as high as the correlation of T^1_{NOR} with NO_2.

Transport of NO to a monitoring site would be expected to decrease the concentration of ozone measured there, due to the atmospheric reaction

$$NO + O_3 \rightarrow NO_2 + O_2$$

Transport of NO, then, may be the reason why local transport from New Orleans does not have a strong net effect on ozone concentration.

The correlation between T_{BTR} and ozone concentration is .102, which is statistically significant at the .10 level. Thus, although this correlation is not high, there is evidence that transport from Baton Rouge increases ozone levels to some extent.

Further work could be done to find a better algebraic combination of variables to represent transport. It is felt, however, that the results above suggest that local transport from the three areas considered does not contribute strongly to ozone concentrations.

TABLE 3-1. Correlation of Pollutant Concentrations With Wind-Direction Variables

	WD NOR	WD BTR	WD HOU
O_3 AV	.027	.039	-.102
NO_x AV	.156	-.067	.159
NO AV	.180	-.080	.175
NO_2 AV	.102	-.049	.103
SO_2 AV	-.051	-.053	.004

TABLE 3-2. Correlation of Pollutant Concentrations With First Set of Local Transport Terms

	T_{NOR}	T_{BTR}	T_{HOU}
O_3 AV	.044	.102	-.083
NO_x AV	.308	-.056	.013
NO AV	.366	-.079	.056
NO_2 AV	.204	-.036	-.016
SO_2 AV	.002	-.034	-.027

TABLE 3-3. Correlation of Pollutant Concentrations With Second Set of Local Transport Terms

	T^1_{NOR}	T^1_{BTR}	T^1_{HOU}
O_3 AV	-.001	.075	-.086
NO_x AV	.360	-.061	.008
NO AV	.436	-.062	.063
NO_2 AV	.235	-.049	-.027
SO_2 AV	-.004	-.027	-.024

3.2 Effects of Wind Direction at Individual Monitoring Sites

In Section 2.0, the area-wide effects of wind direction are discussed. In this section, the effects on ozone concentration at individual monitoring sites are investigated. Sites 8, 10, 14, and 19 were selected for analysis, since more than thirty data points exist for each of them.

In Table 3-4, the correlations between the wind direction variables with average ozone concentration at the four selected sites are presented. It is clear that wind direction effects vary greatly from site to site. For Site 8, for example, the correlations are all .247 or above for the variables WDIR12, WDIR13, WDIR14, and WDIR15 (247.5 to 337.5°, or approximately west to northwest). For the same directions for each of the other three sites, the signs of the correlations are mixed.

Since wind speeds and directions were measured near ground level, however, differences in terrain probably account for some site-to-site differences in wind effects.

The correlations corresponding to WD NOR, WD BTR, and WD NOR are again relatively low in magnitude and have mixed signs. There is some indication, however, that transport from Baton Rouge may affect ozone concentrations at Sites 8, 10, and 19.

A question of interest is, do the 16 wind direction variables, WDIR 1 through WDIR16, explain any part of the day-to-day ozone fluctuations which is not explained by the other variables? That is, do the direction of winds which do not bring pollutants to the sites directly from New Orleans, Baton Rouge, or Houston-Port Arthur relate significantly to the measured ozone concentrations?

TABLE 3-4. Site-Specific Correlations of Average Ozone Concentration with Wind Direction

	Sites Ordered from West to East			
Variable	19	10	8	14
WDIR 1	-.112	.081	-.092	-.039
WDIR 2	.252	-.235	-.126	-.074
WDIR 3	.161	-.006	-.121	-.214
WDIR 4	-.050	.333	-.026	-.237
WDIR 5	-.038	.328	-.140	.177
WDIR 6	-.042	.118	.156	.277
WDIR 7	.161	.177	-.035	.368
WDIR 8	-.054	.171	-.086	.001
WDIR 9	-.210	.011	.109	.187
WDIR10	-.121	-.062	-.150	.198
WDIR11	-.270	.029	-.393	.159
WDIR12	-.053	-.165	.247	-.077
WDIR13	-.135	.006	.303	-.200
WDIR14	.255	-.255	.338	.125
WDIR15	-.138	-.402	.254	-.054
WDIR16	-.227	-.005	.137	-.009
WDNOR	-.120	.003	-.020	.038
WDBTR	.161	.201	.175	-.255
WDHOU	-.050	-.123	.092	-.042

Two regression analyses were performed for each of the four selected sites to address this question. In the first regression, the 16 wind-direction variables were excluded from the set of candidate predictors, and in the second regression, they were included. To illustrate this process, Table 3-5 presents the two regression equations developed for Site 8. Again, stepwise regression analysis was used; thus, if the wind direction variables made a contribution to the equation which was not statistically significant (measurable or repeatable), then the regression algorithm would not have included any of them in the equation.

The results of this analysis are summarized in Table 3-6. In each of the four cases, some of the wind-direction variables entered the equation. The inclusion of these terms increased the correlation from .692 to .792 and decreased the standard error from 18.1 to 16.2 ppb on the average. Thus, the wind-direction terms increase the predictive accuracy to a measurable, but small, extent.

TABLE 3-5. Regression Equations for Average Ozone Concentration at Site 8

Regression Equation Excluding 16 Wind-Direction Variables
Correlation: 0.627
Standard Error: 24.3 ppb

$(O_3$ AV) = 20.8 - 9.00 (WS AV) - 7.70 (WS SD)
- 4.23 (DP AV) + 5.43 (TEMPMX)

Regression Equation Including 16 Wind-Direction Variables
Correlation: 0.748
Standard Error: 21.7 ppb

$(O_3$ AV) = - 22.4 - 6.50 (WS AV) - 2.16 (DP AV)
+ 8.19 (WD BTR) - 6.77 (WDIR 3)
- 11.50 (WDIR 8) - 15.55 (WDIR11)
+ 3.50 (TEMP12)

TABLE 3-6. Comparison of Regression Equations With and Without the 16 Wind-Direction Terms

Site	No. of Data Points	Without 16 Wind-Direction Variables		With 16 Wind-Direction Variables	
		Correlation	Standard Error (ppb)	Correlation	Standard Error (ppb)
8	39	.627	24.3	.748	21.7
10	49	.716	15.0	.780	14.3
14	36	.676	17.4	.818	14.7
19	50	.748	15.5	.823	14.0
AVERAGE	44	.692	18.1	.792	16.2

TABLE 3-6. Comparison of Regression Equations With and Without the 16 Wind-Direction Terms

Site	No. of Data Points	Without 16 Wind-Direction Variables		With 16 Wind-Direction Variables	
		Correlation	Standard Error (ppb)	Correlation	Standard Error (ppb)
8	39	.627	24.3	.748	21.7
10	49	.716	15.0	.780	14.3
14	36	.676	17.4	.818	14.7
19	50	.748	15.5	.823	14.0
AVERAGE	44	.692	18.1	.792	16.2

4.0 DEPENDENCIES AMONG THE METEOROLOGICAL AND SPATIAL VARIABLES

In this section, the effect of dependencies among the meteorological and spatial variables on their relationships with ozone concentration is discussed. To perform this analysis, the partial correlation coefficient has been used (Reference 4). Partial correlation analysis allows the investigation of the relationship between two variables given that one or more other variables are fixed.

An example serves to illustrate the concept. We have seen that wind speed has a correlation of -.198 with average ozone concentration. One might ask the following questions regarding this correlation:

(1) To what extent is the correlation explainable by the fact that wind speed tends to disperse pollutants and, thus, reduce their concentrations?

(2) To what extent is the correlation explainable by the fact that high wind speed is associated with wind persistence (lack of turbulence) and atmospheric stability? When atmospheric stability exists, vertical mixing is limited, which means that little or none of the reservoir of ozone which may exist aloft will be brought to ground level.

To address these questions, one can compute the partial correlation of wind speed with ozone concentration given that wind direction standard deviation, which is a measure of atmospheric instability, is fixed. This partial correlation is -.144, which should be compared to the original correlation of -.198.

The conclusion, then, is as follows:

(1) Wind speed itself, independent of atmospheric instability, is related to ozone concentration. The partial correlation, -.144, indicates, however, that the relationship is not close.

(2) The correlation decreases slightly in magnitude, however, when wind speed standard deviation is fixed. This suggests that part of the original correlation of -.198 is probably explainable in terms of the relationship between wind speed and atmospheric stability.

The conclusions stated above could be drawn with more confidence if the correlations involved were higher.

Along similar lines, the wind direction standard deviation has a correlation of .159 with ozone concentration. When wind speed is fixed, however, the partial correlation between wind speed standard deviation and ozone concentration decreases to .081. Thus, part of the apparent relationship between ozone and the standard deviation is probably explainable in terms of the relationship between the latter and wind speed.

Excluding the nine site variables and the sixteen wind direction variables (WDIR 1, etc.), partial correlation analyses for all variables which have statistically significant correlations with average ozone concentration are presented in Tables 4-1 through 4-7. The "response" referred to in those tables is average ozone concentration.

TABLE 4-1

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN WS AV AND OTHER VARIABLES

CORRELATION WITH RESPONSE -.198 90.00 PCT. CONFIDENCE LIMITS -.292 TO -.100

REGRESSION SLOPE -.2497+01 90.00 PCT. CONFIDENCE LIMITS -.3732+01 TO -.1261+01

VARIABLE	CORRELATION WITH WS AV	PARTIAL CORRELATION OF WS AV WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF WS AV WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
WD SD	-.444	-.144	-.144
D NOR	-.278	-.223	-.169
TEMP12	-.244	-.145	-.137
TEMPMX	-.237	-.153	-.133
TEMPAV	-.201	-.171	-.133
D HOU	.186	-.207	-.151
WD HOU	-.136	-.215	-.169
D BTR	-.129	-.211	-.170

TABLE 4-2

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN TEMPAV AND OTHER VARIABLES

CORRELATION WITH RESPONSE .163 90.00 PCT. CONFIDENCE LIMITS .065 TO .259

REGRESSION SLOPE .7731+00 90.00 PCT. CONFIDENCE LIMITS .3066+00 TO .1240+01

VARIABLE	CORRELATION WITH TEMPAV	PARTIAL CORRELATION OF TEMPAV WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF TEMPAV WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
TEMP12	.966	-.323	-.323
TEMPMX	.961	-.190	-.321
DP AV	.696	.498	-.004
WD SD	.268	.127	-.002
WS AV	-.201	.129	.004
D NOR	-.188	.156	-.038
D BTR	-.129	.155	-.050
D HOU	.122	.161	-.056

TABLE 4-3

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN WD SD AND OTHER VARIABLES

CORRELATION WITH RESPONSE .159 90.00 PCT. CONFIDENCE LIMITS .060 TO .255

REGRESSION SLOPE .1314+01 90.00 PCT. CONFIDENCE LIMITS .5002+00 TO .2128+01

VARIABLE	CORRELATION WITH WS SD	PARTIAL CORRELATION OF WD SD WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF WD SD WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
WS AV	-.444	.081	.081
TEMPMX	.317	.096	.038
TEMP12	.309	.088	.040
TEMPAV	.268	.121	.003
D NOR	.147	.170	.010

TABLE 4-4

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN OP AV AND OTHER VARIABLES

CORRELATION WITH RESPONSE -.261 90.00 PCT. CONFIDENCE LIMITS -.352 TO -.166

REGRESSION SLOPE -.1011+01 90.00 PCT. CONFIDENCE LIMITS -.1384+01 TO -.6375+00

VARIABLE	CORRELATION WITH DP AV	PARTIAL CORRELATION OF DP AV WITH RESPONSE	
		VARIABLE AT LEFT FIXED	VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
TEMPAV	.696	-.529	-.529
TEMP12	.599	-.532	-.448
TEMPMX	.591	-.500	-.445
D NOR	-.287	-.290	-.465
D HOU	.214	-.274	-.472
WD HOU	.116	-.252	-.468

TABLE 4-5

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN WD HOU AND OTHER VARIABLES

CORRELATION WITH RESPONSE -.102 90.00 PCT. CONFIDENCE LIMITS -.200 TO -.002

REGRESSION SLOPE -.2530+01 90.00 PCT. CONFIDENCE LIMITS -.4997+01 TO -.6284-01

VARIABLE	CORRELATION WITH WD HOU	PARTIAL CORRELATION OF WD HOU WITH RESPONSE	
		VARIABLE AT LEFT FIXED	VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
WS AV	-.136	-.133	-.133
DP AV	.116	-.075	-.106

TABLE 4-6

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN TEMPMX AND OTHER VARIABLES

CORRELATION WITH RESPONSE .223 90.00 PCT. CONFIDENCE LIMITS .126 TO .316

REGRESSION SLOPE .1030+01 90.00 PCT. CONFIDENCE LIMITS .5802+00 TO .1480+01

VARIABLE	CORRELATION WITH TEMPMX	PARTIAL CORRELATION OF TEMPMX WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF TEMPMX WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
TEMP12	.964	-.079	-.079
TEMPAV	.961	.242	.073
DP AV	.591	.485	-.044
WD SD	.317	.184	-.037
WS AV	-.237	.185	-.038
D BTR	-.149	.214	-.040
WD BTR	.111	.220	-.056

TABLE 4-7

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN TEMP12 AND OTHER VARIABLES

CORRELATION WITH RESPONSE .252 90.00 PCT. CONFIDENCE LIMITS .157 TO .344

REGRESSION SLOPE .1125+01 90.00 PCT. CONFIDENCE LIMITS .6941+00 TO .1556+01

VARIABLE	CORRELATION WITH TEMP12	PARTIAL CORRELATION OF TEMP12 WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF TEMP12 WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
TEMPAV	.966	.372	.372
TEMPMX	.964	.145	.299
DP AV	.599	.529	.208
WD SD	.309	.217	.211
WS AV	-.244	.215	.203
D BTR	-.155	.243	.199
D NOR	-.136	.247	.209

The most important principles to remember when examining a partial correlation analysis of this type are as follows:

(1) It is reasonable to conclude that a variable has a direct effect on ozone concentration if the following are true:
   (a) The variable's correlation with ozone is not decreased substantially in magnitude as other variables are fixed.
   (b) Physical reasoning can be used to explain how the variable could influence ozone.

Wind speed satisfies these criteria.

(2) Variables whose correlations are drastically reduced in magnitude as other variables are fixed are possibly surrogates for those variables. Physical reasoning must be used to decide which variables have a causative effect on ozone concentration and which are surrogates.

Partial correlation analysis which include the site and wind direction variables are presented in Tables 4-8 through 4-20. In examining these tables, one should remember that any eight of the site variables can serve collectively as a surrogate for the ninth; thus, the partial correlation of any given site variable will always become small when most of the other site variables are fixed. A similar statement applies for the 16 wind direction variables.

Tables 4-4 and 4-15 strongly indicate that the dew point average is not a surrogate for any of the variables analyzed in this section. Evidence is presented in the following section,

TABLE 4-8

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN SITE 8 AND OTHER VARIABLES

CORRELATION WITH RESPONSE .243 90.00 PCT. CONFIDENCE LIMITS .147 TO .335

REGRESSION SLOPE .1787+02 90.00 PCT. CONFIDENCE LIMITS .1076+02 TO .2499+02

VARIABLE	CORRELATION WITH SITE 8	PARTIAL CORRELATION OF SITE 8 WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF SITE 8 WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
D BTR	-.450	.232	.232
WD SD	-.281	.304	.293
SITE19	-.193	.265	.334
SITE10	-.191	.227	.246
WDIR11	.161	.274	.280
SITE14	-.159	.261	.299
WDIR10	.153	.242	.294
D HOU	.147	.242	.309
SITE12	-.138	.223	.266
SITE16	-.132	.233	.022
SITE18	-.132	.235	.001
D NOR	-.122	.239	.008
DP AV	-.111	.223	.009
WD BTR	-.110	.249	.011
WDIR 6	-.101	.252	.007

TABLE 4-9

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN SITE10 AND OTHER VARIABLES

CORRELATION WITH RESPONSE -.118 90.00 PCT. CONFIDENCE LIMITS -.215 TO -.018

REGRESSION SLOPE -.7891+01 90.00 PCT. CONFIDENCE LIMITS -.1453+02 TO -.1249+01

VARIABLE	CORRELATION WITH SITE10	PARTIAL CORRELATION OF SITE10 WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF SITE10 WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
D BTR	-.447	-.173	-.173
WD SD	.312	-.179	-.263
SITE19	-.221	-.103	-.248
SITE 8	-.191	-.075	-.086
SITE14	-.182	-.105	-.049
WD BTR	.170	-.127	-.055
SITE12	-.158	-.153	-.098
WDIR 4	.157	-.129	-.101
SITE16	-.152	-.136	-.220
SITE18	-.152	-.133	-.115
WDIR 2	-.150	-.116	-.113
SITE13	-.113	-.113	.001
D NOR	.110	-.113	.000
WDIR 8	-.101	-.124	.000

TABLE 4-10

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN SITE12 AND OTHER VARIABLES

CORRELATION WITH RESPONSE -.190 90.00 PCT. CONFIDENCE LIMITS -.285 TO -.092

REGRESSION SLOPE -.1611+02 90.00 PCT. CONFIDENCE LIMITS -.2441+02 TO -.7800+01

VARIABLE	CORRELATION WITH SITE12	PARTIAL CORRELATION OF SITE12 WITH RESPONSE	
		VARIABLE AT LEFT FIXED	VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
D NOR	-.352	-.225	-.225
D HOU	.289	-.208	-.236
WDIR 2	.163	-.197	-.238
SITE19	-.160	-.180	-.190
SITE10	-.158	-.213	-.185
WDIR 4	-.139	-.184	-.184
SITE 8	-.138	-.163	-.070
SITE14	-.132	-.182	-.070
DP AV	.119	-.166	-.065
WDIR 7	-.117	-.186	-.062
WDIR14	-.110	-.188	-.053
SITE16	-.110	-.204	-.042
SITE18	-.110	-.202	.001
WDIR15	.108	-.181	.001
WD HOU	.101	-.182	.001

TABLE 4-11

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN SITE16 AND OTHER VARIABLES

CORRELATION WITH RESPONSE -.105 90.00 PCT. CONFIDENCE LIMITS -.203 TO -.005

REGRESSION SLOPES -.9198+01 90.00 PCT. CONFIDENCE LIMITS -.1789+02 TO -.5027+00

VARIABLE	CORRELATION WITH SITE16	PARTIAL CORRELATION OF SITE16 WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF SITE16 WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
WD BTR	-.170	-.100	-.100
WDIR 8	.162	-.098	-.094
SITE19	-.154	-.094	-.084
SITE10	-.152	-.125	-.104
WDIR 4	-.138	-.098	-.099
SITE 8	-.132	-.076	-.046
WD SD	.132	-.129	-.088
SITE14	-.126	-.096	-.055
TEMPAV	.121	-.128	-.063
WDIR 5	-.116	-.095	-.057
SITE12	-.110	-.129	-.105
WD NOR	-.109	-.103	-.100
SITE18	-.105	-.115	-.169
WDIR 6	-.104	-.099	-.162
WDIR16	.100	-.103	-.166

TABLE 4-12

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN WD AV AND OTHER VARIABLES

CORRELATION WITH RESPONSE -.198 90.00 PCT. CONFIDENCE LIMITS -.292 TO -.100

REGRESSION SLOPE -.2497+01 90.00 PCT. CONFIDENCE LIMITS -.3732+01 TO -.1261+01

VARIABLE	CORRELATION WITH WS AV	PARTIAL CORRELATION OF WD AV WITH RESPONSE VARIABLE AT LEFT FIXED	VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
WD SD	-.444	-.144	-.144
WDIR 2	.355	-.220	-.168
WDIR 9	-.318	-.196	-.171
SITE18	-.308	-.237	-.206
WDIR 8	-.302	-.225	-.225
D NOR	-.278	-.223	-.225
WDIR 7	-.252	-.192	-.219
TEMP12	-.244	-.145	-.194
TEMPMX	-.237	-.153	-.191
WDIR11	-.222	-.240	-.230
WDIR 1	.214	-.201	-.231
TEMPAV	-.201	-.171	-.200
WDIR10	-.192	-.196	-.195
D HOU	.186	-.207	-.206
WDIR 6	-.181	-.190	-.218
WDIR16	.175	-.196	-.224
WDIR15	.164	-.185	-.223
SITE14	.148	-.213	-.238
SITE19	.144	-.213	-.250
WDIR14	.143	-.204	-.254
WD HOU	-.136	-.215	-.254
D BTR	-.129	-.211	-.244
WDIR 3	.128	-.202	-.247
WDIR 4	.122	-.207	-.249
SITE17	.108	-.202	-.258

TABLE 4-13

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN TEMPAV AND OTHER VARIABLES

CORRELATION WITH RESPONSE .163 90.00 PCT. CONFIDENCE LIMITS .065 TO .259

REGRESSION SLOPE .7731+00 90.00 PCT. CONFIDENCE LIMITS .3066+00 TO .1240+01

VARIABLE	CORRELATION WITH TEMPAV	PARTIAL CORRELATION OF TEMPAV WITH RESPONSE VARIABLE AT LEFT FIXED	VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
TEMP12	.966	-.323	-.323
TEMPMX	.961	-.190	-.321
DP AV	.696	.498	-.004
WD SD	.268	.127	-.002
WDIR15	-.263	.142	-.003
SITE18	-.226	.148	-.014
WS AV	-.201	.129	-.008
D NOR	-.188	.156	-.033
WDIR 8	.169	.175	-.017
WDIR16	-.143	.161	-.017
SITE17	-.139	.168	-.014
WDIR 7	.137	.159	-.014
WDIR10	.132	.161	-.014
D BTR	-.129	.155	-.047
D HOU	.122	.161	-.047
SITE16	.121	.178	.002
WDIR 3	-.121	.167	-.002
WDIR 6	.117	.157	-.005
WDIR 9	.108	.160	.005

TABLE 4-14

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN WD SD AND OTHER VARIABLES

CORRELATION WITH RESPONSE .159 90.00 PCT. CONFIDENCE LIMITS .060 TO .255

REGRESSION SLOPE .1314+01 90.00 PCT. CONFIDENCE LIMITS .5002+00 TO .2128+01

VARIABLE	CORRELATION WITH WS AV	PARTIAL CORRELATION OF WD SD WITH RESPONSE VARIABLE AT LEFT FIXED	VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
WS AV	-.444	.081	.081
TEMPMX	.317	.096	.038
SITE10	.312	.208	.087
TEMP12	.309	.088	.084
SITE 8	-.281	.245	.166
TEMPAV	.268	.121	.123
WDIR 9	.245	.155	.126
WDIR16	-.148	.157	.128
D NOR	.147	.170	.121
SITE17	-.143	.164	.126
WDIR 8	.137	.168	.130
SITE13	-.133	.168	.144
SITE16	.132	.176	.157
WDIR 1	-.128	.160	.157
WDIR15	-.121	.149	.156
SITE18	.120	.172	.176
WDIR 7	.120	.155	.176
SITE19	-.103	.169	.143

TABLE 4-15

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN DP AV AND OTHER VARIABLES

CORRELATION WITH RESPONSE -.261 90.00 PCT. CONFIDENCE LIMITS -.352 TO -.166

REGRESSION SLOPE -.1011+01 90.00 PCT. CONFIDENCE LIMITS -.1384+01 TO -.6375+00

VARIABLE	CORRELATION WITH DP AV	PARTIAL CORRELATION OF DP AV WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF DP AV WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
TEMPAV	.696	-.529	-.529
TEMP12	.599	-.532	-.448
TEMPMX	.591	-.500	-.445
D NOR	-.287	-.290	-.465
SITE18	-.268	-.297	-.466
D HOU	.214	-.274	-.472
SITE14	.186	-.282	-.494
WDIR12	.172	-.249	-.490
WDIR 3	-.142	-.261	-.489
WDIR16	-.119	-.266	-.490
SITE12	.119	-.245	-.495
WD HOU	.116	-.252	-.498
SITE 8	-.111	-.243	-.494
WDIR15	-.104	-.275	-.496

TABLE 4-16

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN WD HOU AND OTHER VARIABLES

CORRELATION WITH RESPONSE -.102 90.00 PCT. CONFIDENCE LIMITS -.200 TO -.002

REGRESSION SLOPE -.2530+01 90.00 PCT. CONFIDENCE LIMITS -.4997+01 TO -.6284-01

VARIABLE	CORRELATION WITH WD HOU	PARTIAL CORRELATION OF WD HOU WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF WD HOU WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
WDIR12	.795	-.039	-.039
WDIR13	.558	-.090	-.040
WDIR11	.403	-.046	.027
WDIR 3	-.283	-.100	.029
WDIR 4	-.276	-.090	.029
WDIR 5	-.265	-.079	.027
WDIR 2	-.248	-.100	.028
WDIR 6	-.164	-.093	.026
WS AV	-.136	-.133	.025
WDIR 7	-.125	-.097	.026
SITE13	-.120	-.096	.032
DP AV	.116	-.075	.022
WDIR 1	-.106	-.103	.022
SITE12	.101	-.085	.032
WDIR10	.101	-.105	.035
WDIR14	.100	-.105	.027

TABLE 4-17

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN TEMPMX AND OTHER VARIABLES

CORRELATION WITH RESPONSE .223 90.00 PCT. CONFIDENCE LIMITS .126 TO .316

REGRESSION SLOPE .1030+01 90.00 PCT. CONFIDENCE LIMITS .5802+00 TO .1480+01

VARIABLE	CORRELATION WITH TEMPMX	PARTIAL CORRELATION OF TEMPMX WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF TEMPMX WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
TEMP12	.964	-.079	-.079
TEMPAV	.961	.242	.073
DP AV	.591	.485	-.044
WD SD	.317	.184	-.037
WDIR15	-.271	.204	-.038
WS AV	-.237	.185	-.040
SITE18	-.180	.211	-.025
WDIR 8	.159	.235	-.026
SITE17	-.158	.230	-.020
D BTR	-.149	.214	.030
WDIR 7	.134	.219	.030
WDIR16	-.134	.222	.030
WDIR 6	.113	.217	.030
WD BTR	.111	.220	.013
SITE19	.109	.216	.004
WDIR10	.106	.221	.007

TABLE 4-18

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN WDIR11 AND OTHER VARIABLES
CORRELATION WITH RESPONSE -.149 90.00 PCT. CONFIDENCE LIMITS -.245 TO -.050
REGRESSION SLOPE -.4703+01 90.00 PCT. CONFIDENCE LIMITS -.7816+01 TO -.1590+01

VARIABLE	CORRELATION WITH WDIR11	PARTIAL CORRELATION OF WDIR11 WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF WDIR11 WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
WD HOU	.403	-.119	-.119
WDIR10	.348	-.169	-.140
WDIR 3	-.269	-.149	-.141
WS AV	-.222	-.202	-.173
WDIR 4	-.217	-.140	-.171
WDIR 2	-.192	-.148	-.171
WDIR12	.168	-.135	-.183
SITE 8	.161	-.197	-.215
WDIR 5	-.144	-.137	-.214
WDIR 6	-.131	-.142	-.211
WDIR13	.106	-.145	-.201
WDIR14	-.106	-.147	-.152

TABLE 4-19

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN WDIR15 AND OTHER VARIABLES

CORRELATION WITH RESPONSE -.102 90.00 PCT. CONFIDENCE LIMITS -.200 TO -.003

REGRESSION SLOPE -.2559+01 90.00 PCT. CONFIDENCE LIMITS -.5042+01 TO -.7536-01

VARIABLE	CORRELATION WITH WDIR15	PARTIAL CORRELATION OF WDIR15 WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF WDIR15 WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
WDIR14	.411	-.125	-.125
TEMP12	-.303	-.028	-.051
TEMPMX	-.271	-.045	-.050
TEMPAV	-.263	-.062	-.032
WDIR 3	-.228	-.101	-.031
WDIR 4	-.206	-.093	-.026
WDIR 5	-.204	-.084	-.018
WDIR 2	-.181	-.100	-.011
WDIR 6	-.178	-.092	-.007
WS AV	.164	-.072	.024
WDIR16	.139	-.100	.026
WDIR 7	-.136	-.097	.029
WD SD	-.121	-.085	.029
WDIR 9	-.111	-.099	.034
SITE12	.108	-.084	.055
DP AV	-.104	-.135	.071

TABLE 4-20

CORRELATIONS SIGNIFICANT AT 90.0 PCT. LEVEL BETWEEN TEMP12 AND OTHER VARIABLES

CORRELATION WITH RESPONSE .252 90.00 PCT. CONFIDENCE LIMITS .157 TO .344

REGRESSION SLOPE .1125+01 90.00 PCT. CONFIDENCE LIMITS .6941+00 TO .1556+01

VARIABLE	CORRELATION WITH TEMP12	PARTIAL CORRELATION OF TEMP12 WITH RESPONSE VARIABLE AT LEFT FIXED	PARTIAL CORRELATION OF TEMP12 WITH RESPONSE VARIABLE AT LEFT AND THOSE ABOVE IT FIXED
TEMPAV	.966	.372	.372
TEMPMX	.964	.145	.299
DP AV	.599	.529	.208
WD SD	.309	.217	.211
WDIR15	-.303	.234	.214
WS AV	-.244	.215	.207
SITE18	-.202	.241	.198
WDIR16	-.159	.252	.198
D BTR	-.155	.243	.222
WDIR 7	.155	.248	.222
WDIR 8	.145	.263	.208
WDIR 6	.143	.246	.211
D NOR	-.136	.247	.211
WDIR10	.128	.251	.210
SITE17	-.116	.257	.203
WDIR 9	.116	.250	.204

however, which indicates that dew point and temperature serve as surrogates for solar radiation.

5.0 CORRELATION OF SOLAR RADIATION WITH OTHER VARIABLES

Solar radiation was measured at only three sites, 10, 12, and 19, while ambient temperature was measured at all sites. For this reason, temperature was employed in the detailed statistical study, but solar radiation was not. In this section, however, a brief analysis of the relationships of solar radiation with the other variables is presented. One hundred and eighteen data points were employed.

Table 5-1 presents the correlations between average ozone concentrations and several other variables. Being based on a subset of the data sample of 273 points discussed in the preceding sections, some of the correlations in Table 5-1 differ to a moderate extent from those presented in Table 2-2.

Table 5-1 shows that the average solar radiation correlates .392 with O_3 AV; this is somewhat higher than the correlations of maximum radiation, TEMPAV, TEMPMX, or DP AV with ozone.

The correlation between TEMPAV and average radiation is .417, while the correlation between the maximum temperature and the maximum radiation is .565. Thus, solar radiation and temperature are definitely related, but the relationship is not an extremely close one.

Moreover, when solar radiation is fixed, the correlations of both average temperature and average dew point with ozone concentration are reduced in magnitude somewhat. The correlation of TEMPAV and O_3 AV based on the 118 data points is .28; the partial correlation with solar radiation fixed is .14, half as large. The corresponding correlation and partial correlation for DP AV are -.23 and -0.18, respectively; the decrease in the case is only .05.

TABLE 5-1. Correlations of Several Variables With Average Ozone Concentration

Variable	Correlation with O_3 AV
Av. Solar Radiation*	.392
Max. Solar Radiation*	.294
TEMPAV	.277
TEMPMX	.310
DP AV	-.232

*As is the case with the other meteorological variables, these values represent conditions from 8 a.m. to 4 p.m. The "maximum solar radiation" is the maximum hourly average.

The indication is that the relationships that TEMPAV and DP AV appear to have with O_3 AV are probably due at least in part to the relationships of these two variables with solar radiation.

In Table 5-2, an analysis is presented which is designed to test the extent to which TEMPAV and DP AV can serve as surrogates for radiation in a predicting equation. It is shown that the partial correlations of both average and maximum radiation with ozone become negative given that TEMPAV and DP AV are fixed. On a very intuitive basis, this can be interpreted to mean that TEMPAV and DP AV in combination are capable of explaining essentially all of the positive relationship between radiation and ozone concentration. This does not negate the point made earlier that radiation plays a direct role in ozone formation through photochemical reactions. The point is simply that other variables can substitute for radiation in a predictive equation.

Along the same lines, Table 5-3 presents the multiple correlation and standard error as successive terms are entered into an equation to predict O_3 AV. Average solar radiation was forced to enter last, so that its <u>incremental</u> predictive value could be assessed.

Although it is statistically significant, the incremental contribution of solar radiation to the predictive accuracy is small. Its inclusion results in an increase from .712 to .727 in correlation and a decrease from 16.8 to 16.5 ppb in the standard error. The final equation is as follows:

$$(O_3\ AV) = -6.8 + 13.5\ (SITE19) + 4.43\ (TEMPAV) - 3.99\ (DP\ AV) - .419\ (R)$$

TABLE 5-2. Partial Correlations of Solar Radiation With Average Ozone Concentration

	Variables Fixed	Partial Correlation with Average Ozone Concentration of: Av. Solar Radiation	Max. Solar Radiation
1.	None	.392	.294
2.	TEMPAV	.316	.188
3.	TEMPAV, DP AV	-.209	-.302

TABLE 5-3. Multiple Correlations of Sets of Variables with Average Ozone Concentration

	Set of Variables	Multiple Correlation with Average Ozone Concentration	Standard Error (ppb)
1.	TEMPAV	.277	22.8
2.	TEMPAV, DP AV	.654	18.0
3.	TEMPAV, DP AV, SITE19	.712	16.8
4.	TEMPAV, DP AV, SITE19, Av. Solar Radiation	.727	16.5

where R is average solar radiation from 8 a.m. until 4 p.m. in langleys per hour.

The fact that solar radiation has a negative coefficient is due to the presence of the other terms in the model which correlate with radiation. If an equation were developed including only radiation, its coefficient would be positive.

6.0 OZONE CORRELATIONS AMONG SITES

In this section, a small study of the relationships among the ozone concentrations at the different sites is discussed. The hourly average ozone concentrations were employed, and only hours for which data were available from all sites were used. Thus, there is no danger that variations among the computed correlations is due to an irrelevant influence such as difference in hours of day or dates for which data were used.

It was possible, however, to analyze only data for 42 hours on this basis. The hours of day from 8 a.m. to 8 p.m. are well represented, but all data points were taken from September 5 through September 9, 1976. There was rain throughout the area during this period, but the temperatures were not depressed.

The ozone means and standard deviations are presented in Table 6-1.

The correlations between hourly ozone concentrations at pairs of sites are presented in Tables 6-2 and 6-3. In Table 6-2, the correlations are ordered according to geographical position, and in Table 6-3, they are ordered according to NO_x and hydrocarbon emission density in the county (see Figure 2-3).

The site-to-site correlations are relatively high, ranging from .629 to .932. This indicates a definite relationship among the sites with respect to ozone concentration by hour.

Table 6-2 does not reveal any strong geographical patterns in the site-to-site correlations. By reading across the row for Site 18, however, there does seem to be a decreasing correlation between Site 8 and other sites as the distance between the two sites being correlated increases. Site 18 is the farthest west and the most isolated of the sites.

TABLE 6-1. Ozone Means and Standard Deviations by Site for 42 Hours Occurring on September 5-9, 1976

SITES ORDERED FROM WEST (TOP) TO EAST	MEAN OZONE CONCENTRATION (ppb)	STANDARD DEVIATION OF OZONE CONCENTRATION (ppb)
18	55.3	20.1
19	55.3	19.4
10	40.6	14.4
16	48.0	15.6
8	63.6	29.4
17	44.8	19.5
12	31.7	11.5
13	75.3	26.4
14	71.8	16.7

TABLE 6-2. Site-by-Site Correlations Between Hourly Ozone Concentrations, Ordered by Site Location

Monitoring Sites Ordered From South to North	Monitoring Sites Ordered from West to East								
	18	19	10	16	8	17	12	13	14
10	.857	.914	1.000	.860	.790	.807	.877	.738	.878
8	.734	.799	.790	.748	1.000	.893	.773	.671	.768
19	.926	1.000	.914	.932	.799	.753	.772	.763	.811
18	1.000	.926	.857	.862	.734	.713	.640	.753	.719
14	.719	.811	.878	.835	.768	.718	.891	.767	1.000
12	.640	.772	.877	.772	.773	.771	1.000	.626	.891
17	.713	.753	.807	.704	.893	1.000	.771	.639	.718
16	.862	.932	.860	1.000	.748	.704	.772	.782	.835
13	.753	.763	.738	.782	.671	.639	.626	1.000	.767

TABLE 6-3. Site-by-Site Correlations Between Hourly Ozone Concentrations, Ordered by County NO_x and Hydrocarbon Emission Densities

Site \ Site	NO_x<24 HC<19			NO_x<24 19<HC<44	NO_x>57 HC<19	NO_x>57 HC>44			
	10	13	16	19	8	12	14	17	18
18	.857	.753	.862	.926	.734	.640	.719	.713	1.000
17	.807	.639	.704	.753	.892	.771	.718	1.000	.713
14	.878	.767	.835	.811	.768	.891	1.000	.718	.719
12	.877	.626	.772	.772	.773	1.000	.891	.771	.640
8	.790	.671	.748	.799	1.000	.773	.768	.893	.734
19	.914	.763	.932	1.000	.799	.772	.811	.753	.926
16	.860	.782	1.000	.932	.748	.772	.835	.704	.862
13	.738	1.000	.782	.763	.671	.626	.767	.639	.753
10	1.000	.738	.860	.914	.790	.877	.878	.807	.857

HC - Hydrocarbon emission density by parish.

NO_x - NO_x emission density by parish.

Both are in tons/sq.mi./yr.

Table 6-3 indicates that the correlations are fairly uniform with regard to the relative emission densities of the areas containing the two sites being correlated. The emissions inventory data, however, are not believed to be highly accurate and do not reflect emissions from oil fields and other possible rural sources.

Due to the small time period considered, the results of this section should be considered tentative. While time limitations do not allow further work during the current project, additional study of the relationships among ozone concentrations at different sites would be beneficial.

7.0 SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

This report discusses a statistical study of ozone concentrations in southern Louisiana from July 17 to October 17, 1976. Data collected at nine monitoring sites were analyzed.

It has been shown that average daytime ozone concentration at different sites has a statistically significant (that is, measurable or repeatable) positive correlation with

(1) wind direction standard deviation,
(2) ambient temperature, and
(3) solar radiation.

Ozone concentration correlates negatively with

(1) average wind speed, and
(2) average dew point.

Of these five, average solar radiation has the highest correlation, .392, with average ozone concentration.

The statistical evidence indicates that dew point and ambient temperature in combination can serve as surrogates for solar radiation; that is, they are capable of replacing radiation in an equation to predict average ozone concentration. This in no way, however, denies the direct causative role of radiation in the formation of ozone through atmospheric photochemical reactions.

There is evidence, moreover, that the relationships that both ambient temperature and dew point appear to have with ozone concentration are due at least in part to the correlations these two variables have with radiation. This is based on a partial correlation analysis in which it was shown that fixing

the value of solar radiation lessened the strength of the relationships of the two variables with ozone concentration somewhat.

It is difficult to separate the effects of wind speed and wind direction standard deviation, which have a correlation of -.444. Strong winds disperse pollutants of all kinds, and thus, reduce their concentrations.

Atmospheric instability, however, is associated with vertical mixing, which decreases the concentrations of most pollutants. It can increase the concentration of ozone in the morning though by bringing large amounts of ozone from aloft to ground level. Strong winds are negatively related and large wind direction standard deviations are positively related to atmospheric instability.

There is not an extremely strong area-wide dependence of ozone concentration on wind direction. Ozone tends to be slightly lower, however, when the wind is from the west and northwest.

There is evidence of site-to-site differences in the effect of wind direction on ozone concentration. Since wind speeds and direction were measured near ground level, however, differences in terrain probably account for some site-to-site differences in wind effects.

Transport from the New Orleans and Houston-Port Arthur areas apparently has very little effect on ozone concentration. Transport from New Orleans, however, does increase the concentration of NO_x. It is suspected that transport of NO may explain why local transport does not have a strong net effect on ozone levels; the presence of NO reduces the concentration of ozone

through atmospheric chemical reactions. Transport from Baton Rouge increases ozone levels to a small, but statistically significant, extent.[1]

A regression equation was developed to predict average ozone concentrations in terms of the meteorological and spatial variables. The equation has a correlation of .732, and a standard error of 18.0 ppb. The standard error can be thought of as the standard deviation of the differences between the measured concentrations of ozone and the corresponding predicted values. The correlation and standard error quantify the closeness of the relationship between (1) the meteorological and spatial variables in combination and (2) ozone concentration.

A small study was made of the correlations between pairs of sites in hourly average ozone concentrations during the day. The site-to-site correlations were fairly high, ranging from .629 to .932. The correlation has a slight tendency to decrease as the distance between the two sites being correlated increases. The amount of NO_x and hydrocarbon emissions surrounding the two sites apparently has little effect on the correlation. The available emissions data are not believed to be highly accurate, however, and do not reflect emissions from all fields and other possible rural sources.

The results of the hourly average ozone correlation study should be considered tentative, since only the period between September 5 and September 9, 1976, was considered. This restriction, however, allowed the inclusion of only hours when data from all nine sites were available.

---

[1]A detailed case study of three computer-generated trajectories indicated stronger evidence of transport from Baton Rouge. (Reference 3)

The authors feel that additional research into the behavior of ozone as a function of time of day at the different sites would be beneficial. This research could be done with the same data base used in this study.

The objectives of this research would be to answer the following physical questions by using multivariate statistical methods:

(1) What is the shape of the curve of ozone concentration versus time of day?

(2) How does the curve change from site to site?

(3) How do the various meteorological variables affect the curve?

Once the curves were developed, they could be used to compute the average ozone concentration between any given hours of the day or to compute the maximum daily ozone concentration and its time of occurrence for a given site and set of meteorological conditions. Perhaps more importantly, an expanded understanding of the influences on ozone concentration would be gained.

The use of computer-generated air trajectories to explain variations in ozone levels is another possible future area of study.

REFERENCES

1. Draper, N. R., and H. Smith, Applied Regression Analysis, John Wiley & Sons, Inc., New York, 1966.

2. Karl, T. R., and G. A. DeMarrais, "Meteorological Conditions Conducive to High Levels of Ozone," Proceedings of the International Conference on Photochemical Oxidant Pollution and Its Control, January 1977.

3. Lambeth, Bryan W., "Meteorological Case Studies of Extreme Summer Ozone Events in Southern Louisiana," Technical Note, Radian Corporation, Austin, Texas, December 1977.

4. Morrison, Donald F., Multivariate Statistical Methods, McGraw-Hill Book Company, New York, 1967.

5. Ott, Sun-Tsui, and Walter A. Lyons, "Further Evidence of Long Range Photochemical Oxidant Transport Inferred from Acoustic Sounder Data," Joint Conference on Applications of Air Pollution Meteorology, American Meteorological Society, Salt Lake City, Utah, November 29-December 2, 1977.

6. Wark, Kenneth, and Cecil F. Warner, Air Pollution, Its Origin and Control, IEP, New York, 1976.

APPENDIX A
ADDITIONAL INFORMATION REGARDING THE MULTIPLE REGRESSION MODEL FOR AVERAGE OZONE CONCENTRATION

This appendix supplements the discussion in Section 2.4 about the multiple regression model for average ozone concentration. The information here sheds additional light on the model, but is not absolutely required for an understanding of the basic observations given in the text. This discussion is included, however, for the benefit of readers who are interested in additional statistical detail regarding the regression model. Statistical terms are used more freely here than in the text.

Assumptions Made in the Multiple Regression Analysis

Before performing a regression analysis, it is necessary to specify the algebraic form of the equation to be used. Regression, then, can be used to compute the coefficients of the preselected terms. If the wrong assumptions are made regarding the form of the equation, then the regression model may not agree well with the data.

Stepwise regression analysis (Reference 1) was used to develop the equation presented here. This method greatly reduces the likelihood of specifying the wrong functional form, since the method can select from a possibly very large set of candidate predictor variables only those which contribute significantly to the accuracy of the equation. Thus, in this analysis, all of the spatial and meteorological variables listed in Table 2-1 were candidate predictors. The following equation, also presented in Section 2.4, was developed:

$$
\begin{aligned}
(O_3\ AV) = \\
24.4 &+ 17.8\ (SITE\ 8) + 21.4\ (SITE\ 13) + 19.8\ (SITE\ 14) \\
&+ 24.7\ (SITE17) + 20.8\ (SITE19) \\
&- 3.34\ (WS\ AV) - 2.60\ (DP\ AV) \\
&+ 0.0542\ (D\ HOU) + 1.98\ (WDIR\ 2) \\
&- 2.30\ (WDIR\ 8) + 2.46\ (WDIR10) \\
&- 6.02\ (WDIR11) + 2.89\ (WDIR14) \\
&+ 2.58\ (TEMP12)
\end{aligned}
$$

The correlation coefficient for the equation is .732, and the standard error is 18.0 ppb. The standard error can be thought of as the standard deviation of the differences between the predicted and measured values. The equation was developed from 273 data points.

The terms in the model were initially assumed to be linear, unless a physical reason existed for including products, quotients, or powers of the basic variables. Interaction terms designed to account for local transport were considered, and none improved the predictive accuracy. Local transport is discussed in Section 3.

Various plots of measured minus predicted ozone concentrations versus predictor variables were produced, and no evidence was seen that the inclusion of a nonlinear function of any one of the variables would improve the equation. The plots provided a check on the validity of using linear terms in the equation.

Thus, it is felt that a reasonable effort has been made to include the appropriate terms in the model.

Certain additional mathematical assumptions are required if regression techniques are to be strictly valid; these assumptions are virtually never all satisfied exactly in applied scientific and engineering work.

Possibly the most important of these assumptions is that the dependent variable (ozone concentration in this case) may have a random error, but that the independent variables must be error-free. If this assumption is not satisfied, the probable effect is that the regression coefficients will be smaller in magnitude, the standard error will be larger, and the multiple correlation will be smaller.

This assumption is not strictly satisfied in the study presented here. The meteorological variables are undeniably affected by measurement error.

As is discussed in the text, however, the basic objective of the regression analysis was simply to quantify the strength of the relationship between (1) the spatial and measured meteorological variables as a group and (2) the measured ozone concentration. Development of a predictive model as an end in itself was not an important objective of the study. It is felt that the multiple correlation and standard error given above provide a reasonable indication of the strength of the relationship of interest.

Additional Statistics Regarding Multiple Regression Analysis

In this section, additional statistics are presented regarding the multiple regression model. The statistical terms employed here are widely used and are discussed in books such as Reference 1.

Table A-1 presents the analysis of variance for the multiple regression. The primary conclusion from this table is that there is no reasonable doublt that the meteorological and spatical variables as a group have a measurable (repeatable) relationship with average ozone concentration. This follows from the fact that the F-ratio is 21.3, which is statistically significant at well above the 1% level.

TABLE A-1
ANALYSIS OF VARIANCE FOR MULTIPLE REGRESSION ANALYSIS

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Ratio
Regression	14	96626	6901	21.3
Residual	258	83499	323	

In Table A-2, the regression coefficients, the standard errors of the regression coefficients, and the F-statistics for the individual terms are given. The standard errors in this table can be thought of as the standard deviations of the random errors in the coefficients.

The standard errors are useful in assessing whether serious multi-collinearities exist among the terms in the model. If so, standard errors which are large compared to the regression

coefficients themselves can result. It can be seen from the table that this situation does not exist here. This is not surprising, since, when two terms are very highly correlated with each other, the stepwise algorithm usually does not include both of them in the model.

Moreover, the analysis of the relationships among the individual predictor variables and ozone concentration was made in terms of correlations and partial correlations rather than in terms of the multiple regression coefficients anyway. This analysis is presented in Sections 2.2, 3, 4, and 5. Thus, the study does not hinge on the meaningfulness of the regression coefficients, which are difficult to interpret in many real-world multivariate problems.

The F-statistics given in Table A-2 are the basis for including the given terms in the equation. All of these values are 2.0 or higher, which means that all given terms are significant at the .25 level; all but the smallest two are significant at the .10 level. Moreover, when each of the listed variables was added to the equation, the overall standard error for the equation was reduced; this is another indication that each of the terms included makes a significant contribution to the predictive accuracy.

For each of the terms excluded from the equation, the F-statistic was no larger than 1.2. This value is not statistically significant at the .25 level.

TABLE A-2

STATISTICS RELATED TO MULTIPLE REGRESSION ANALYSIS

Variable	Coefficient	Standard Error	F Ratio
SITE 8	17.8	3.5	25.5
SITE 13	21.4	5.4	15.5
SITE 14	19.8	4.3	20.7
SITE 17	24.7	9.4	6.9
SITE 19	20.8	4.1	26.3
WS AV	- 3.34	.69	23.4
DP AV	- 2.60	.22	133.9
D HOU	0.0542	.0325	2.8
WDIR 2	1.98	1.06	3.5
WDIR 8	- 2.30	1.61	2.0
WDIR 10	2.46	1.61	2.4
WDIR 11	- 6.02	1.48	16.6
WDIR 14	2.89	1.03	7.9
TEMP 12	2.58	.26	100.4